VRIJE UNIVERSITEIT AMSTERDAM

VU Research Portal

Discrete choice models with availability constraints

de Palma, A.; Rouwendal, J.

1990

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)
de Palma, A., & Rouwendal, J. (1990). *Discrete choice models with availability constraints*. (Serie Research Memoranda; No. 1990-54). Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam.

Faculteit der Economische Wetenschappen en Econometrie

Serie Research Memoranda

Discrete choice models with availability constraints

André de Palma
Jan Rouwendal

Research Memorandum 1990-54
November 1990

vrije Universiteit *amsterdam*

Discrete Choice Models with Availability Constraints*

André de Palma** and Jan Rouwendal***

november 1990

JEL Calssification numbers : 210, 920, 930.
Keywords : Availability Constraints, Discrete Choice Models, Dutch Housing Market, Search Behaviour, Strict Utility Models.

* We wish to thank M. Ben-Akiva and A. de Vos for helpful comments.

** Dept. of Civil Engineering, Northwestern University, Evanston (Ill.), USA.

*** Dept. of Economics, Free University, Amsterdam, The Netherlands.

The first author's research was supported by the National Science Association, award no. SES-8909730, the second author's research by the Netherlands Organization for Scientific Research (NWO), award no. 46-250.

Abstract

In this paper we study choice behaviour in situations where one out of a finite number of alternatives has to be chosen and the choice set is unknown. The relevance of this framework for actual choice behaviour is discussed and a general model is developed. It is shown that this model can be viewed as the outcome of a process of sequential search. The model is first presented in general form, while several special cases are dealt with later on. It is shown that, apart from information about the probabilities that alternatives will be available, the informational requirements of the model are - under general assumptions - comparable to those of conventional discrete choice models. The paper concludes with an application to search behaviour in the Dutch housing market.

1 Introduction

Discrete choice models describe how an individual facing a set of exclusive alternatives selects one of them. A large part of the literature is devoted to probabilistic choice models (for a review see e.g. Anderson, de Palma and Thisse, 1990). The source of the uncertainty in these models can occur at two levels : (i) the model maker is incompletely informed about the behaviour of the individuals or (ii) the behaviour of the individuals themselves is inherently probabilistic. In the former case each individual is acting according to specific but idiosyncratic rules and perceived variables; the model maker is only partially informed about the specific behaviour of each individual and the best he can do is to model individual behaviour up to a probability distribution. In the latter case each individual is acting according to a probabilistic rule, i.e. an individual facing the same situation two consecutive times may select different choices.

These two polar cases give rise to probabilistic choice models which are formally equivalent, at least if one uses them - as will be the case in this paper - to describe the behaviour of an 'average' individual in some 'homogeneous' population in one period. 'Homogeneity' of the population means that all individuals are identical from a statistical point of view and have independent preferences.

We will be concerned here with the introduction of another type of uncertainty in the standard disrete choice models described above. Instead of assuming that the set of alternatives is given, we will assume that some alternatives may not be available for some individuals. For example, a household searching for a new dwelling may find out that some types of vacant dwellings are not available anymore at the time they are selected. Alternatively, a consumer patronizing a shop may find out that it does not have the item she is looking for. Standard discrete choice models do not incorporate this kind of uncertainty. For this reason we have to consider an alternative framework by assuming there is a given (super)set of possible choice sets and that each possible choice set has a given probability of being relevant. This means that individuals face a probabilistic availability constraint. The probabilistic process that governs the availability of alternatives is assumed to be independent from the other sources of uncertainty that have been mentioned above. In other words : the ranking of the various alternatives is independent of their availability. Since there is little reason to assume such a dependency, this seems to be a reasonable assumption. Our approach excludes, therefore, among others, any search strategy where alternatives are visited sequentially as a function of their expected availability.

The formulation presented in this paper is closely related to the analysis of Ben-Akiva (1977), who proposed a model of choice set generation (see also the discussion in de Palma and Lefèvre, 1981, and Swait and Ben-Akiva, 1987a and b). Ben Akiva and his coauthors have developed a model in which random constraints are the basis of the choice set formation process and extend therefore the captivity models as described by Gaudry and Wills (1979). For mode choice models, for example, the availability of a mode (car or bus) depends on both deterministic and probabilistic constraints. The car is, for example, not available if the commuter has no driving licence (except if he carpools); however, the distance to the nearest transit acts more as a probabilistic constraint. Moreover, the lack of complete information at the analyst's point of view and the presence of a very large number of potential alternatives justify the use of probabilistic choice set generation (see, for example, Wermuth, 1978, Ben-Akiva et al. 1984, and Train, McFadden and Ben-Akiva, 1984). Note that, although the general idea of probabilistic choice set generation developed in this article is very similar to that developed by Swait and Ben-Akiva (1978a and b), the interpretation of the model is different. The availability of alternatives depends in this article on the '<u>supply-side</u>' (e.g., a dwelling unit is or is not available for rent) and not, as is assumed in the articles mentioned above, on the 'demand side' (as in the case of a budget constraint which restricts the set of availble alternatives). As a consequence the choice set generating process and the choice process will be assumed to be independent in what follows. This property will allow us to simplify the expressions for the choice probabilities. In this article we study the mathematical implications of models with probabilistic supply-side constraints and derive simple probabilistic discrete choice expressions assuming specific behavioral preference patterns.

The organisation of the paper is as follows. We will first present the model in general terms in section 2. In section 3 we give an alternative interpretation. Two examples of explicit models are given in sections 4 and 5, while sections 6 and 7 demonstrate the empirical usefulness of the model by means of an application to search behaviour in the Dutch housing market. Section 8 concludes the paper.

2 Presentation of the basic model

We consider a situation in which an actor faces a set of N+1 potential alternatives. He does not know which subset of these alternatives will actually be available to him. One alternative, to be denoted by means of an index 0, is assumed to be always included in the subset that is relevant to the actor, so that we are guaranteed that the actor has always at least one option available. This alternative may be identified with the continuation of the original situation by the actor.[1]

The actual choice set of the actor will be denoted as S and consists of the alternative 0 and a subset of the set of alternatives $C' = \{1,\dots,N\}$:

$$S = \{0\} \cup T, \qquad (1)$$

with $T \in \Omega(C')$, where $\Omega(C')$ denotes the power set of C'. We will refer to $C = \{0\} \cup C'$ as the potential choice set and to S as the actual choice set. The set of sets S that satisfy (1) will be denoted as Ψ.[2]

The probability that an individual selects alternative n when facing an actual choice set S will be denoted as $\pi(n,S)$. The characteristics of the alternatives and the prices are assumed to be exogenously determined. The probability π may be described by an arbitrary discrete choice model that satisfies the two assumptions stated below. Note that the uncertainty with respect to the availability of alternatives does not play a role for the determination of $\pi(n,S)$, since S is known.

The probability that n will be selected if it does not belong to the set S is, of course, zero. Formally :

Assumption 1

$$\pi(n,S) = 0, \quad \text{if } n \notin S. \qquad (2)$$

Moreover, we will assume that the probability that alternative n will be chosen will never increase if the choice set S is extended. Formally :

Assumption 2

$$\pi(n,S^1) \leq \pi(n,S^2), \quad \text{if } S^2 \subset S^1 \text{ and } n \in S^2. \qquad (3)$$

These two assumptions are intuitively satisfied and will be used later on in this article.

The overall probability $p(n,C)$ that an alternative n will be chosen within the potential choice set equals C can be determined as follows :

$$p(n,C) = \sum_{S \in \Psi(C)} Q(S,C).\pi(n,S), \qquad (4)$$

where $Q(S,C)$ denotes the probability that S is the actual choice set when the potential choice set equals C.

Equation (4) provides the relation between the standard choice probabilitities $\pi(n,S)$, defined for a given choice set S of available alternatives, and the choice probabilities $p(n,C)$, defined over a potential choice set C. The situation which is usually analyzed by discrete choice models corresponds to the specific case where $Q(C,C) = 1$ and $Q(S,C) = 0$ for $S \neq C$; in this case $p(n,C)=\pi(n,C)$.

One may wonder how the choice probabilities $p(n,C)$ differ from the corresponding probabilities $\pi(n,C)$. In order to compare the two probabilities, it must be observed that two effects of the uncertainty can be distinguished. In the first place there is the possibility that a particular potential alternative n will not be actually available, i.e. $n \notin S$ for some $S \in \Psi(C)$. In this case $\pi(n,S)$ equals 0. This has a decreasing effect on the choice probability $p(n,C)$ as compared to $\pi(n,C)$, cf. assumption 1 and equation (4). In the second place, however, it may occur that alternative n is available, while other potential alternatives are not, a situation which may be expected to increase the probability $p(n,C)$ that alternative n will be chosen, cf. assumption 2. The total effect of the uncertainty is therefore ambiguous.

The two effects may be illustrated by rewriting (4) as :

$$p(n,C) = \pi(n,C). \sum_{S \in \Psi(C)} Q(S,C).[\pi(n,S)/\pi(n,C)]. \qquad (5)$$

Define the probability q_n that alternative n will be contained in the choice set as :

$$q_n = \sum_{S \in \Psi_n(C)} Q(S,C), \qquad n=1\ldots N, \tag{6}$$

where $\Psi_n(C)$ is the subset of $\Psi(C)$ containing only those actual choice sets S for which $n \in S$, as defined above. The variables q_n are of course nonnegative and at most equal to 1.

Now return to (5) and consider two examples. First, assume that $q_n < 1$ and $\pi(n,S) = \pi(n,C)$ whenever $n \in S$. It follows immediately that $p(n,C) = q_n . \pi(n,C) < \pi(n,C)$. Second, assume that $q_n = 1$ and $\pi(n,S) > \pi(n,C)$ for all $S \neq C$ with $n \in S$. It follows that $p(n,S) > \pi(n,S)$, except for the special case in which $Q(S,C)=0$ for all $S \in \Psi_n(C)$, $S \neq C$.

These two examples show that the impact of availability constraints on the choice probabilities is in general ambiguous. There is one special case, however, that occurs when $n=0$. Since alternative 0 is always available, only the second effect exists and $p(0,C) \geq \pi(0,C)$, as expected.

Although the total effect of the availability constraints is ambiguous, we can define the intervals in which the probabilities $p(n,C)$ should be :

Proposition 1

$$q_n . \pi(n,C) \leq p(n,C) \leq q_n, \qquad n=1\ldots N. \tag{7}$$

Proof: Alternative n can only be chosen when it is available (assumption 1). This implies immediately : $p(n,C) \leq q_n$. It follows from assumption 2 that $\pi(n,S) \geq \pi(n,C)$ whenever $n \in S$. Substitution in equation (4) then gives the second inequality : $q_n . \pi(n,C) \leq p(n,C)$. □

Finally, it should be noted that the average effect of the availability constraints on the probabilities that the alternatives 1,...,N will be selected satisfies the following condition :

$$\sum_{n \geq 1} p(n,C) \leq \sum_{n \geq 1} \pi(n,C). \tag{8}$$

The reason is that the probability that alternative 0 will be chosen can only increase as a consequence of the availability constraints.

3 An alternative interpretation

Until now we have interpreted our model as referring to an actor who first determines his actual choice set and then chooses one of the alternatives which are available. In the present section we will show that an alternative interpretation can be given to the model. Consider the following situation. An actor knows the potential choice set, but not the actual one. He therefore acts in the following way. He chooses the alternative that he prefers most, say n, then looks whether it is available. When this is the case the process stops. If n is not available he goes on to choose the alternative in the choice set C/{n} which he prefers most, looks whether it is available and continues iteratively until an available alternative has been found. Given that alternative 0 is always available, we can be sure that always an alternative will be selected by this procedure. The significance of this second interpretation is that, in many cases, it seems to be much more realistic as a description of actual choice behaviour.

If the individual preferences and the availability of alternatives are (as we have assumed) independent stochastic processes, the two interpretations are equivalent. To prove this, it suffices to consider a realization of the stochastic process defining the preference ordering of an individual and the availability of alternatives.

The choice behaviour of an individual is governed by two independent random processes. The first of these determines the ranking $\rho(C)$ of the various alternatives by the individual, while the second determines the actual choice set S(C) which is relevant for him. Given the determinateness of ρ and S for a specific individual, his choice behaviour can be described in two alternative ways :

(i) the most preferred feasible choice is the most preferred choice in the set of feasible choices S; note that, once the feasible choice set is determined, the choice set reduces to an usual discrete choice problem. This is the way we have interpreted our model in the previous section.

(ii) the best feasible alternative can also be be defined as the first available choice in the ranking of ρ. This is the alternative interpretation introduced in the present section.

These two ways to describe the best feasible alternative selected by the individual clearly give rise to the same outcome. This conclusion holds for any realization of the random processes that determine the ranking and the set of available alternatives so that the two descriptions are equivalent.

In other words, the probability that a given alternative, say n, is selected by the individual can alternatively be described as the sum of the probabilities that a particular choice set S containing n is available multiplied by the probability that the best alternative in S is n (see equation 4), or as the probability that the first available alternative in the ranking is n.

This can be formalized in the following way. Let $\rho(S)$ denote a ranking of the elements of S and let $\rho_n(S)$ denote such a ranking when n is the most preferred alternative in S. The relation between the two interpretations can be written down as follows :

$$\mathrm{Prob}(\rho_n(S)) = \sum_{\rho(C)\,|\,\rho_n(S)\in\rho(C)} \mathrm{Prob}(\rho(C)), \qquad (9)$$

where the notation $\rho_n(S)\in\rho(C)$ means that the ranking of the elements of S in $\rho(C)$ is the same as in $\rho_n(S)$. With the second interpretation the choice probabilities p(n,C) can be written as :

$$p(n,C) = \sum_{S\in\Psi(C)} Q(S,C) \,.\, \sum_{\rho_n(S)} \sum_{\rho(C)\,|\,\rho_n(S)\in\rho(C)} \mathrm{Prob}(\rho(C)), \qquad (10)$$
$$n=0,\ldots,N.$$

Now observe that :

$$\pi(n,S) = \sum_{\rho_n(S)} \mathrm{Prob}(\rho_n(S)), \quad n\in S. \qquad (11)$$

After substitution of (11) in (10) we arrive at equation (4). This proves rigourously the equivalence of the two interpretations.

4 A simple example

In order to be able to specify a discrete choice model with availability constraints, one needs to have information about the probabilities that the various possible choice sets will be realized and about choice probabilities of the actor under the various actual choice set. This implies, generally speaking, that a large amount of information about the individual actor and about the environment in which he has to make his decisions should be available. It is therefore useful to show that there are cases in which these informational requirements can be reduced to modest proportions. For this purpose we will - in the present and next section - deal with two examples of specific discrete choice models according to the second interpretation proposed.

Probably the most simple example of a discrete choice model with availability constraints arises when it is assumed that individuals either realize their first choice or choose alternative 0. Such behaviour may be relevant when a time constraint allows the searchers to look for dwellings of one type only. This situation arises for example on the housing market, where the actors in general do not consider all units in the market, but limit their search to a few of them. Such behaviour may imply that the choice probabilities $\pi(n,S)$ are always equal to $\pi(n,C)$ when $n \in S$. The observable probabilities $p(n,C)$ can therefore be determined as :

$$p^{1}(n,C) = \begin{cases} q_n.\pi(n,C) & \text{when } n=1,\ldots,N \\ \pi(0,C)+\sum\limits_{n'\notin S}(1-q_{n'}).\pi(n',C) & \text{when } n=0 \end{cases} . \qquad (12)$$

It is clear from equation (12) that in this - admittedly simple - case the informational requirements are modest. When one is able to specify the standard choice probabilities and one has information about the probabilities that the various alternatives will be available, the model can be specified easily.

5 Independence of Irrelevant Alternatives

It would of course be nice if the conclusion reached in the last paragraph could be shown to be valid for more general types of choice behaviour as well. In order to do this, we will make two additional assumptions. The first postulates a kind of independence in the way the choice sets are generated, the second does something comparable for the way the rankings of the alternatives are generated.

Assumption 3 The probability that alternative n is contained in the actual choice set S is independent of the probability that any other alternative n' ($n' \neq n$) is contained in S.

This assumption allows us to conclude immediately that, for example, the probability that alternatives n and n' are included in an arbitrary choice set and alternative n'' is not included equals $q_n.q_{n'}.(1-q_{n''})$. It follows from the assumption that the probability Q(S,C) that the actual choice set is S equals :

$$Q(S,C) = \prod_{n\in S} q_n . \prod_{n'\notin S} (1-q_{n'}). \qquad (13)$$

Assumption 3 thus gives rise to a simple expression of the realization probabilities Q(S,C).

Assumption 4 The ratio of two choice probabilities $\pi(n,S)$ and $\pi(n',S)$, $n,n' \in S$, is independent of the other alternatives contained in S.

The consequences of this second assumption may be explored by noting that the ratio $\pi(n,S)/\pi(n',S)$ is independent of S, and can therefore be written as :

$$\frac{\pi(n,S)}{\pi(n',S)} = \frac{\alpha_n}{\alpha_{n'}}, \quad n,n'\in S, \qquad (14)$$

where the α_n's are positive parameters, one of which can be chosen arbitrarily. It is easy to deduce the following expression for the choice probabilities from this formula :

$$\pi(n,S) = \frac{\alpha_n}{\sum_{n'\in S} \alpha_{n'}}, \quad n\in S. \qquad (15)$$

The best-known example of a discrete choice model satisfying equation (15) is the multinomial logit model, but it should be noted that other models (e.g. the probit model when the variance-covariance matrix of its error terms is diagonal) may also satisfy it.

Assumption 4 is the well-known independence of irrelevant alternatives (IIA) assumption, which has been criticized on theoretical grounds, but is, nevertheless, very often used in empirical work. Models characterized by equation (15) have been studied extensively in mathematical psychology and have been named strict utility models by Block and Marschak (1960). Note that assumptions 1 and 2 are satisfied by strict utility models.

One may wonder whether assumption 4 implies that the probabilities p(n,C) also satisfy the IIA assumption. A glimpse at equation 4 shows that this will, in general, not be the case.

There is an implication of assumption 4 that deserves further notice. After dividing both the numerator and the denominator of the RHS of equation (12) by $\sum_{n''\in C}\alpha_{n''}$, we can write :

$$\pi(n,S) = \frac{\pi(n,C)}{\sum_{n'\in S}\pi(n',C)}. \tag{16}$$

The significance of this expression is that it relates choice probabilities that refer to a choice set S to choice probabilities that refer to the complete choice set C.

We now substitute (16), together with (13), in (4) and obtain :

$$p(n,C) = q_n.\pi(n,C).\sum_{S\in\Omega'_n(C)}[\prod_{\substack{n'\in S\\ n\neq n'}} q_{n'}.\prod_{n''\notin S}(1-q_{n''})] \;/\; [\sum_{n'\in S}\pi(n',C)]. \tag{17}$$

In this equation only choice probabilities that are relevant when C is the actual choice set occur, so that it can be concluded that we have, once again, succeeded in reducing the informational requirements about the behaviour of the actor to the usual ones.

Comparison of equation (17) with (12) shows that a complicated third term has been added to the simple formulation $q_n.\pi(n,C)$. It represents the increase in the probability that alternative n will be chosen that results from the fact that the actor may, in the present situation, decide to choose alternative n as a second-best (or third-best, fourth-best, etc.) choice, when the most preferred alternatives are not available to him. Such behaviour reminds one of satisficing behaviour.[3]

6 Towards an empirical application

In section 4 we have considered a model in which searchers looked whether their first best alternative was available and chose alternative 0 when it was not. In the present section we will study extensions of that model in which searchers may visit a fixed number of alternatives, which may be larger than one. For a motivation of such models we refer again to housing market search. Consider a situation in which a household is looking for a dwelling. It regards a particular type of dwelling as its first best choice, but realizes that such a dwelling may not be available. Two options exist in the case of non-availability : look for a dwelling of the second best choice or stop searching. When the first alternative is chosen it may be the case that a dwelling of the second-best type is also not available. The household then has to decide whether it will look for a third-best choice or stops searching. The process can in principle continue until all existing dwelling types have been considered. However, in practice one may expect that a household will stop the searching process earlier, since the returns of continuing the search process are considered to be too low to make it worthwhile.

In the present section we will consider models in which the actor has decided to stop the search process, by choosing alternative 0, after a finite number, i ($1 \leq i \leq N$), of alternatives have been visited.[4] In section 4 we considered the case in which i equals 1, here we will generalize to larger values.

We first recall the following result. Let $\rho = (n_1, \ldots, n_{N+1})$ denote a particular ranking of the N+1 choice alternatives. We have :

Proposition 2 If the choice probabilitites satisfy a strict utility model, then the ranking probabilities are given by :

$$\mathrm{Prob}(\rho) = \pi(n_1, C).\pi(n_2, C/\{n_1\}). \ \ldots \ .\pi(n_N, \{n_N, n_{N+1}\}). \qquad (18)$$

This proposition has been proved by Block and Marschak [1960, p. 109] and was quoted and discussed by Luce and Suppes [1965, p. 354].[5] Using the expression (16) for the choice probabilities, we obtain easily an expression for the probability of any ranking.

Note that expression (18) implies that the events "n_1 is the most preferred alternative in C", "n_2 is the most preferred alternative in $C/\{n_1\}$", etc. can be treated as independent events. In general this is not true. However, for strict utility models it is valid.

Now consider the situation in which i=2, i.e., the searching actors who discover that their first-best alternative is not available will switch to their second-best alternative. When both alternatives are unavailable, search will be stopped. The probability that alternative n will be chosen can in this case be determined as :

$$p^2(n,C) = q_n.\pi(n,C) + \sum_{\substack{n'=1\\ n'\neq n}}^{N} q_n.(1-q_{n'}).\mathrm{Prob}(n_1=n',n_2=n), \qquad (19)$$

$$n=1\ldots N.$$

The first term on the right-hand-side (RHS) of this equation is identical to the one given in (12) for $n\neq 0$. The second term is new and refers to actors who have realized alternative n as their second-best choice. The expression behind the summation sign is the probability that alternative n is the second-best choice for someone who had alternative n' as his first choice, multiplied by the probability $(1-q_{n'})$ that alternative n' is not available and by the probability q_n that alternative n is available. In order to find the total number of second-best choices for alternative n we have to sum over all possible first-best choices n', excluding n.

The problem is now to find an expression for the probability that n' is ranked first and n second. It follows from proposition 2 that :

$$\mathrm{Prob}(n_1=n',n_2=n) = \pi(n',C).\pi(n,C/\{n'\}). \qquad (20)$$

Using (16), (20) can be rewritten as :

$$\mathrm{Prob}(n_1=n',n_2=n) = \pi(n',C).\frac{\pi(n,C)}{1-\pi(n',C)}, \qquad (21)$$

and after substitution of this result in (19) we finally find :

$$p^2(n,C) = q_n.\pi(n,C)\ \{\ 1 + \sum_{\substack{n'=1\\ n'\neq n}}^{N} (1-q_{n'}).\frac{\pi(n',C)}{1-\pi(n',C)}\ \}, \qquad (22)$$

$$n=1\ldots N.$$

We now move on to the situation in which i=3. It may happen then that also third-best choices are realized by impatient searchers. This implies that we have to add a third term to the right-hand-side of (19). The equation becomes :

$$p^3(n,C) = q_n.\pi(n,C) + \sum_{\substack{n'=1\\ n'\neq n}}^{N} q_n.(1-q_{n'}).\mathrm{Prob}(n_1=n',n_2=n) + \sum_{\substack{n'=1\\ n'\neq n}}^{N} \sum_{\substack{n''=1\\ n''\neq n\\ n''\neq n'}}^{N} q_n.(1-q_{n'}).(1-q_{n''}).\mathrm{Prob}(n_1=n',n_2=n'',n_3=n), \qquad n=1\ldots N. \tag{23}$$

The interpretation of the third term on the RHS is analogous to that of the second one. The probability that n' and n'' are not available, whereas n is available is $q_n.(1-q_{n'}).(1-q_{n''})$. Using again proposition 2, the probability that n' is ranked first, n'' second and n third can be written as :

$$\mathrm{Prob}(n_1=n',n_2=n'',n_3=n) = \pi(n',C).\pi(n'',C/\{n'\}).\pi(n,C/\{n',n''\}), \tag{24}$$

or :

$$\mathrm{Prob}(n_1=n',n_2=n'',n_3=n) = \pi(n',C).\frac{\pi(n'',C)}{1-\pi(n',C)}.\frac{\pi(n,C)}{1-\pi(n',C)-\pi(n'',C)}. \tag{25}$$

Therefore, the probability that n is selected is in this case equal to :

$$p^3(n,C) = q_n.\pi(n,C).\Big(1 + \sum_{\substack{n'=1\\ n'\neq n}}^{N} (1-q_{n'}).\frac{\pi(n',C)}{1-\pi(n',C)}.\Big[1 + \sum_{\substack{n''=1\\ n''\neq n\\ n''\neq n}}^{N} (1-q_{n''}) .\frac{\pi(n'',C)}{1-\pi(n',C)-\pi(n'',C)} \Big] \Big), \qquad n=1\ldots N. \tag{26}$$

It should be clear by now that we can extend our model again in order to deal with the possibility that fourth-best choices are made, etc. These generalisations are conceptually straightforward. Instead of writing down all these equations, we will confine ourselves to the statement of :

Proposition 3

$$p^{i+1}(n,C) \geq p^{i}(n,C) \qquad (27)$$

for i=1,...,N-1, but :

$$[p^{i+2}(n,C)-p^{i+1}(n,C)] < [p^{i+1}(n,C)-p^{i}(n,C)]. \qquad (28)$$

Proposition 3, which is proved in appendix 1, says that the probability of selecting an alternative increases (as expected) as the number of alternatives that can be visited during the period under consideration increases, but that the marginal increment in this probability decreases with the number of alternatives that can be visited. This can be interpreted as decreasing returns to search. The proposition allows us to see the various stages (when the individual searches for 1,2,3,.. alternatives) as successive approximations. This is the procedure that will be followed in the empirical application of the next section.

In that section we will use observations on the number of realized moves for the determination of the availability probabilities. We will then determine the model estimates of the flow of actors leaving the various subpopulations of searchers for a particular alternative and compare these with figures observed in an alternative way. The correlation between the two figures will be regarded as an indication of the validity of the model.

We will use a sequential approach by first employing the model of section 4, then allowing for second-best choices, then also for third-best choices, etc. It will of course be expected that the introduction of switching possibilities will, at least initially, improve the correlation between computed and actual numbers of leavers.

7 Search behaviour in the Dutch housing market

We will analyse search behaviour in the heavily regulated rented segment of the Dutch housing market. Rent control was introduced in the market in the Second World War and has been covering almost all rented dwellings ever since. Although many governments intended to return to a (more or less) free market all attempts to do so failed thus far. The main reason for this failure is that rent control is regarded by many as a part of income policy. Rent increases, which are inevitable for a return to free-market circumstances are therefore hard to realize politically. Since the attempts of the government to raise construction levels of new dwellings to a levels, by means of subsidies, have not been able to remove the shortages, excess-demand has been the normal situation on the rented part of the housing market ever since the Second World War.

The dwellings that become available are distributed by local authorities, who give priority to households that are judged to be especially in need of a(nother) dwelling. However, the rules that are used (if they really exist) are certainly not uniform and are far from being clear. The outcome of the process for the individual searching household is highly unpredictable and the resulting situation may perhaps be best described as one of uncertainty about the actual possibilities.

Before presenting the results we give some details about the data. We used the Dutch Housing Needs Survey (WoningBehoefte Onderzoek, WBO) of 1981, a 1% sample of all Dutch households consisting of more than 60,000 respondents. This database contains information about the willingness to move to another type of dwelling of households, as well as on past mobility (for all households in the sample) and on the realized duration of search (for households willing to move). We focussed on moves between rented dwellings in the same region. (See DGVH [1982a] for further information.)

The stock of rented dwellings has been classified in 16 types. We used a dichotomy in apartments and single-family dwellings and made further divisions which refer to the number of rooms and the rent to be paid. In order to take into account the regional differences in housing market situation we divided the Netherlands into four regions. Details are given in appendix 2.

Since searching households were asked to indicate some characteristics of their most desired dwelling, we were able to determine the number of households searching for each of the dwelling types considered. The ratio of this number and the total number of searchers has been interpreted as an estimate of the choice probability $\pi(n,C)$, to be denoted as $\hat{\pi}(n,C)$, where C denotes the set of sixteen dwelling types that have been distinguished. This means that we have regarded the total population of searchers as a group of statistically independent individuals, an assumption which can only be defended as a convenient first approximation.[6]

In order to get an estimate of p(n,C), the probability that a searching household will move to a dwelling of type n, we employed the information about past mobility contained in the sample. The average yearly number of households moving to a given type of dwelling in the 4 years preceding 1981 was regarded as a good indicator of the number of households in the sample that would be making such a move in 1981. So the ratio of the average number of realized moves in a dwelling of type n in the preceding years and the total number of searching households can be interpreted as an estimate of $p(n,C)$, to be denoted as $\hat{p}(n,C)$. Note that this procedure can only be justified if the market situation is constant over time.[7]

The values of $\hat{p}(n,C)$ and $\hat{\pi}(n,C)$ can be used to arrive at estimates of the availability probabilities q_n. For the simple model of section 4 this is easy : we simply take the ratio of $\hat{p}(n,C)$ and $\hat{\pi}(n,C)$ (see equation (12) for $n \neq 0$). For the more complicated models in which more than one alternative may be visited in one period this is more difficult. We used an iterative procedure, started with arbitrary values of the q_n's and substituted these, together with the $\hat{\pi}(n,C)$'s into the model equations (for i=2 equations (22), for i+3 equations (26), etc.). The resulting values give the model prediction of the $p(n,C)$'s, which will be denoted as $\tilde{\tilde{p}}(n,C)$. If $\tilde{\tilde{p}}(n,C)$ was larger than $\hat{p}(n,C)$, the value if q_n was decreased, in the reverse case q_n was increased.[8] This procedure allowed us to make the sum of squared differences $\sum_n (\tilde{\tilde{p}}(n,C) - \hat{p}(n,C))^2$ arbitrary small.[9] The final values of the q_n's will be denoted as $\hat{q}_n$.

The estimates of the availability probabilities and the unconstrained choice probabilities can be used to predict the number of searchers that intended to move to a dwelling of a particular type that leave this population in one period. These numbers will be denoted as ℓ_n. The households that leave a population of searchers may either move to a dwelling of their most desired type, but may also have left for a dwelling of another type (second-best choice, third-best choice, etc.). The relevant equations for i, the number of alternatives to be considered, equal to 1,2 and 3 are :

$$\ell^1_n = q_n.\pi(n,C).b, \tag{27}$$

$$\ell^2_n = (q_n.\pi(n,C) + (1-q_n) \sum_{n' \neq n} q_{n'}.\frac{\pi(n',C)}{1-\pi(n,C)}).b, \tag{28}$$

$$\ell_n^3 = \{q_n.\pi(n,C) + (1-q_n)\sum_{n'\neq n} q_{n'}.\frac{\pi(n',C)}{1-\pi(n,C)} +$$

$$+ (1-q_n)\sum_{n'\neq n}(1-q_{n'}).\frac{\pi(n',C)}{1-\pi(n,C)}\sum_{\substack{n''\neq n\\ n''\neq n'}} q_{n''}.\frac{\pi(n'',C)}{1-\pi(n,C)-\pi(n',C)}\}.b, \tag{29}$$

where b is the total number of searching households. The equations for higher values of i are analogous. Substitution of the estimates $\hat{q}_n$ and $\hat{\pi}(n,C)$ in these equations leads to the predicted values of the numbers of leavers, to be denoted $\tilde{\tilde{\ell}}_n^i$. As a test of model performance these predicted values of the numbers of leavers will be compared with estimates of these numbers which are obtained in a completely different way.

The Housing Needs Survey contains information about the realized duration of search of people intending to move to another dwelling. This information can be used to estimate of the yearly number of households leaving a subpopulation of searchers. We employed a simple duration-of-search model to determine the rate of leaving the subpopulation of households searching for a particular type of dwelling. This enabled us - again on the basis of an assumption of stationarity - to determine the yearly number of households leaving these subpopulations. Details about this procedure and its results are given in appendix 3. The estimates of the numbers of leavers will be denoted as $\hat{\ell}_n$.

In table 1 we have listed our estimates of the numbers of households that realize a move, intend to make a move and leave the population of searching households, respectively. It is clear from this table that the number of realized moves is in almost all cases greater than the number of intended moves. The ratios between these numbers give an estimate of the availability probabilities, which was shown above to be valid when the simple model of section 4 is relevant. However, if we adopt that model we expect the numbers of people who realized a move to be equal to the numbers of people leaving the population of searching households. The first row of table 2 shows that these figures are indeed correlated to some extent, but they are hardly impressive.

There are various possibilities to interpret this result. Households can leave the population of searchers by moving to a dwelling of the desired type,

Table 1 Movers and searchers on the housing market

type	realized moves region 1	2	3	4	intended moves region 1	2	3	4	households leaving the population of searchers region 1	2	3	4
1	21.5	104.8	76.0	35.5	69.0	265.0	182.0	101.0	43.9	103.2	76.5	54.4
2	7.5	26.8	8.0	11.3	40.0	176.0	130.0	71.0	20.2	76.9	54.1	36.5
3	4.3	17.0	9.0	2.3	11.0	66.0	50.0	36.0	3.2	36.0	19.7	14.8
4	16.5	40.8	14.3	17.3	14.0	60.0	39.0	25.0	6.7	20.4	17.0	7.9
5	30.0	77.3	18.3	35.0	65.0	303.0	204.0	97.0	36.5	125.0	82.6	45.7
6	10.3	56.0	23.8	27.0	19.0	138.0	116.0	61.0	12.6	72.6	53.0	29.9
7	29.3	61.5	21.0	31.8	45.0	120.0	117.0	45.0	23.0	45.8	43.4	16.6
8	14.0	46.0	28.0	18.8	16.0	84.0	81.0	29.0	12.1	40.1	36.8	16.1
9	4.3	16.5	72.5	7.0	10.0	29.0	87.0	13.0	6.8	14.4	49.4	6.9
10	10.5	38.3	57.8	17.0	13.0	86.0	128.0	28.0	9.5	42.7	52.9	14.8
11	9.3	33.8	69.0	7.5	14.0	59.0	135.0	12.0	6.3	26.0	59.4	6.8
12	8.8	27.0	60.3	17.5	30.0	153.0	290.0	57.0	15.6	76.7	119.4	27.0
13	2.5	14.8	33.3	5.3	8.0	40.0	85.0	24.0	5.6	17.5	38.6	9.2
14	11.3	26.0	41.0	9.3	4.0	19.0	71.0	7.0	2.0	8.2	21.6	2.8
15	12.5	69.3	84.8	27.0	23.0	80.0	225.0	30.0	12.4	34.5	92.7	13.6
16	3.5	26.5	51.5	10.5	4.0	53.0	102.0	17.0	2.0	28.1	46.9	7.6

by moving to another type of dwelling, or by disappointedly stopping the search process altogether. It was shown above that the first possibility in itself does not give a very good explanation of the figures listed in table 1. The importance of the second possibility can be investigated by adopting the more complicated model that have been introduced in section 6. This will be done in what follows. The third possibility will not be analyzed in detail. We will introduce it explicitly in the last paragraphs of the present section in order to deal with the differences between estimated and predicted numbers of leavers that remain after the possibilities of two and more visits in one period have been introduced.

We have adopted the models introduced in section 6, estimated the availability probabilities and used the result to predict the numbers of households leaving the populations of searchers. These predictions were

Table 2 Comparison of computed and actual numbers of households leaving the population of searchers*

number of visits (i)	correlation coefficients by region 1	2	3	4
1	.745	.658	.531	.674
2	.962	.934	.914	.944
3	.979	.964	.971	.973
4	.982	.973	.984	.981
5	.984	.977	.988	.984
:	:	:	:	:
16	.984	.984	.986	.987

* The figures are correlation coefficients.

compared with the estimates given in table 1. The correlation coefficients are shown in table 2.

It can be inferred from this table that the introduction of the possibility of a second-best choice gives rise to a remarkable improvement in the correlation between predicted and estimated numbers of households leaving the population of searchers. This improvement is largest for region 3, where the excess demands are especially high. Introduction of the possibility of third, fourth and fifth visits gives rise to further increases in the correlation coefficients.

The figures in Table 2 show that the introduction of the possibility of visits to alternatives that were not on the top of the ranking of the individual household gives rise to an important improvement of model performance. It should be observed that this improvement occurs even though we have made strong assumptions on the behaviour of individual searchers (all households were considered to be statistically indistinguishable). It may be useful to stress, once again, that our estimates of the number of leavers were derived completely independent from the model predictions and that there is no a priori reason why the correlation coefficients should be steadily increasing with the number of switches and approach the upper bound 1 so closely.

The typical pattern one would expect for the figures in table 2 is an increase in the correlation coefficients for a relatively small number of visits, followed by a decrease. The number of visits corresponding with the maximum correlation should give an indication of the actual number of reconsideration of choices as a result of the experienced impossibility of realizing initial choices. The figures listed in table 2 do not show this pattern. The correlation coefficients show a rapid increase to a level close that corresponding with 16 visits. Since the correlation coefficients in this case are rather high, it is almost impossible to find a maximum, followed by a decrease.

When the number of visits equals 16 the model says that a move will take place whenever at least one alternative is available. This implies that the ratio of the number of leavers and the number of searchers will be equal for all the subpopulations of searchers. The correlations between the predicted and actual numbers of leavers will therefore in these case be equal to the correlations between the numbers of intended moves and the numbers of actual leavers, which turned out to be very high. The figures are shown in the last row of table 2.

The fact that the correlation coefficients of the numbers of intended moves and the numbers of households leaving the subpopulations of searchers are so close to one would seem to indicate that many households accept any available alternative, i.e. that the actual number of visits may be as high as 16. It suggests, moreover, that the number of leavers from a given subpopulation is (almost) independent of the realization probability of their indicated first-best alternative. This conjecture was affirmed by least squares regressions.[10]

What table 2 shows, therefore, is that the apparent lack of a relation between numbers of realized moves and numbers of leavers can be explained by allowing the searching households to reconsider their initial choices 2 or 3 times, making the total number of visits equal to 3 or 4. Availability constraints are clearly able to weaken the relation between stated preferences, referring to first-best choices, and observed behaviour as measured by numbers of leaving households.

The predicted numbers of households leaving the various populations of searchers increase with the number of switches, although at a rapidly decreasing rate. This means that large changes in the correlation coefficients should not be expected to occur after the introduction of the possibility of six, seven or more visits. For this reason we stopped the exercise after considering the possibility of up to five visits.

The predicted numbers of leavers are, with a few exceptions, lower than the estimated numbers of leavers from table 1. For regions 1,2 and 4 these numbers are on the average 10 % lower than those of table 1, for region 3 the difference amounts to 23 %. These deviations can be explained by existence of a 'seventeenth alternative', viz., stopping the search without realizing a move. This possibility has not been considered explicitly hitherto. It is quite conceivable that the number of disappointed searchers is much higher in the Rimcity, where housing market problems are concentrated, than in the other parts of the Netherlands.

The number of households that realize their first-best choice, as predicted by the model, is steadily decreasing with the number of alternatives that may be visited. When only first-best choices are allowed, all movers realize the alternative they ranked highest. When also second-best choices are allowed, the percentage of them realizing their first best choice drops to the range 53-62 %. Introduction of the possibility of third-best choices changes this range to 35-45 %, introduction of fourth-best choices to 26-37 %, introduction of fifth-best choices to 20-32 %. The implication of the model, therefore, is that many households will move to another type of dwelling than they indicated as their first-best choice. The fact that the increase in the correlation coefficient becomes smaller after the introduction of the second-best choices, while the percentage of moving households still drops significantly, makes it difficult to say whether the actual percentage is 25 or 45. The reason is that the model predicts relatively large numbers of households realizing their third-, fourth- or fifth-best choices.

It must be concluded that our empirical exercise, although it is quite favorable for our model, does not make clear how many visits should be allowed in the model.

8 Conclusion

In the foregoing discussion a general model for discrete choice with availability constraints has been developed. Some special cases have been formulated and an application to the analysis of search behaviour in the Dutch housing market has been provided. Other applications could refer to the labour market or to shopping behaviour.

Notes

1 Alternative 0 should be identified with what happens to our actor when, unfortunately, none of the alternatives whose availability is uncertain are actually available.

2 Note that $\Psi(C)$ is neither the power set of C nor the power set of $\{0\}\cup C$.

3 It may also have to do with time preference. The actor has to compare the loss of the higher utility, associated with giving up the search for a more preferred alternative, which is currently unavailable, with the somewhat lower utility of choosing an available, but less preferred alternative. This trade-off occurs only in cases where switches between alternatives are costly and will not be made in every period. It is therefore more important when the model refers to search on the housing market than when it is concerned with shopping behaviour.

4 The determination of the number of visits may be based on a comparison of the costs of additional search with its expected benefits. When the searchers switches to alternatives which are ranked lower, the expected benefits decrease, while the costs remain the same. Moreover, the searcher may decide to take another chance on high-ranked alternatives in the next period, while a move to a less preferred alternative in the present period makes this alternative less attractive because of transaction costs that have to be paid.

5 We are grateful to Moshe Ben-Akiva to provide us with this reference.

6 It is of course to be expected that preferences depend on household size, income, etc.

7 Some evidence in favour of this assumption can be found in Scholten [1988].

8 Since we have $\partial\pi_n/\partial q_n \geq 0$ and $\partial\pi_n/\partial q_{n'} \leq 0$, $n' \neq n$, we can be sure that our procedure works into the right direction when the changes in the q_n's are small.

9 We have not investigated the statistical properties of our estimation method. It can be interpreted as minimization of the sum of squared deviances of equations in which the q_n's appear in a non-linear way.

10 We used the predicted number of leavers as our dependent variable, which was regressed on the number of intended moves and the realization probability. A singnificant positive effect of the latter variable could only be found for region 3, which is the Dutch Rimcity where housing market problems are concentrated.

References

Anderson, S., A. de Palma and J. Thisse [1990] Analysis of Differentiated Markets with Discrete Choice Models, MIT Press, Cambridge (Ma.),in progress.

Ben-Akiva, M. [1977] Choice Models with Simple Choice Set Generating Processes, working paper, Dept. of Civil Engineering, MIT, Cambridge (Ma).

Ben-Akiva, M.,M. Bergman, A. Daly and R. Ramaswany [1984] Modelling Route Choice Behaviour, Proc. of the 9th Int. Conf. on Traffic and Transportation Theory, VNU, The Netherlands.

Block, H.D. and J. Marchak [1960] Random Orderings and Stochastic Theories of Response, in I. Olkin (ed.) Contributions to Probability and Statistics, Stanford University Press, Stanford (Cal.), 97-132.

de Palma, A., and C. Lefèvre [1981] A Probabilistic Search Model, Journal of Mathematical Sociology, 8, 43-60.

DGVH [1982a] Woningbehoefte Onderzoek 1981 (housing Needs Survey 1981), Ministerie van VROM, Den Haag.

DGVH [1982b] Woningbouwprogrammering : Een Nadere Uitwerking (Dwelling Construction Control: An Elaboration), part 2, Ministerie van VROM, Den Haag.

Gaudry, M. and M. Wills [1979] The Dogit Model, Transportation Research, 13B, 105-111.

Luce, R.D. and P. Suppes [1965] Preference, Utility and Subjective Probability, in R.D. Luce, R.R. Bush and E. Galenter (eds.) Handbook of Mathematical Psychology, III, Wiley, New York, 249-410.

Rouwendal, J. [1989] Choice and Allocation Models for the Housing Market, Kluwer, Dordrecht.

Rouwendal, J [1990] Housing Choice and Search Behaviour in a Disequilibrated Market; An Exploratory Analysis, research paper, Free University, Amsterdam.

Scholten, H.J. [1988] Verhuisprocessen op de Nederlandse Woningmarkt (Residential Mobility in the Dutch Housing Market), GI, Utrecht.

Swait, J. and M. Ben-Akiva [1987a] Incorporating Random Constraints in Discrete Models of Choice Set Generation, Transportation Research, 21B, 91-102.

Swait, J. and M. Ben-Akiva [1987b] Empirical Test of a Constrained Choice Discrete Choice Model : Mode Choice in Sao Paulo, Brazil, Transportation Research, 21B, 103-115.

Train, K.E., D. McFadden and M. Ben-Akiva [1987] The Demand for Local Telephone Services : A Fully Discrete Choice Model of Residential Calling Patterns and Service Choices, Rand Journal of Economics, 18, 109-123.

Wermuth, M. [1978] Structure and Calibration of a Behavioral and Attitudinal Binary Choice Model Between Public Transportation and Private Car, PTRC Meeting, Univ. of Warwick, UK.

Appendix 1 Proof of proposition 3

We can write down the following general expression for p^i (using a slightly different notation) :

$$p^i(n,C) = q_n.\mathrm{Prob}(n_1=n) +$$
$$+ \sum_{n(1)\neq n} q_n.(1-q_{n(1)}).\mathrm{Prob}(n_1=n(1),n_2=n) +$$
$$\cdots$$
$$+ \sum_{n(1)\neq n} \sum_{\substack{n(2)\neq n\\ n(2)\neq n(1)}} \cdots \sum_{\substack{n(i-1)\neq n\\ n(i-1)\neq n(1)\\ \vdots \\ n(i-1)\neq n(i-2)}} q_n.[\prod_{j=1}^{i-1}(1-q_{n(j)}))].\mathrm{Prob}(n_1=n(1),..,n_i=n).$$
$$i=2,\ldots,N.$$

The last term of the RHS of this equation equals p^i-p^{i-1}. Since this term is nonnegative, it follows that the first part of the proposition is proved.

We have also to show that the final term is non-increasing in i. To do this, it suffices to show that :

$$\mathrm{Prob}(n_1=n(1),,\ldots,n_i=n) \geq \sum_{n(i)} (1-q_{n(i)}).\mathrm{Prob}(n_1=n(1),,\ldots,n_i=n(i),n_{i+1}=n),$$

where the summation takes place under the proper restrictions. Using proposition 2 we write :

$$\mathrm{Prob}(n_1=n(1),,\ldots,n_i=n) =$$
$$= \pi(n(1),C)..\pi(n(i-1),C/\{n(1),..,n(i-2)\}).\pi(n,c/\{n(1),..,n(i-1)\})$$

$$= \pi(n(1),C)..\pi(n(i-1),C/\{n(1),..,n(i-2)\}).\frac{\pi(n,C)}{1-\sum_{j=1}^{i-1}\pi(n(j),C)}$$

$$= \pi(n(1),C)..\pi(n(i-1),C/\{n(1),..,n(i-2)\}).\sum_{n(i)}\frac{\pi(n(i),C)}{1-\sum_{j=1}^{i-1}\pi(n(j),C)}.\frac{\pi(n,C)}{1-\sum_{j=1}^{i-1}\pi(n(j),C)}$$

$$> \pi(n(1),C)..\pi(n(i-1),C/\{n(1),..,n(i-2)\}).\sum_{n(i)} \frac{\pi(n(i),C)}{1-\sum_{j=1}^{i-1}\pi(n(j),C)}.\frac{\pi(n,C)}{1-\sum_{j=1}^{i}\pi(n(j),C)}$$

$$= \sum_{n(i)} \mathrm{Prob}(n_1=n(1),,...,n_i=n(i),n_{i+1}=n)$$

$$\geq \sum_{n(i)} (1-q_{n(i)}).\mathrm{Prob}(n_1=n(1),,...,n_i=n(i),n_{i+1}=n).$$

The first step is simply the equation given in proposition 2; the second follows from equation (13); to understand the third step, observe that $\sum_{n(i)} \pi(n(i),C) = 1-\sum_{j=1}^{i-1}\pi(n(j),C)$ when the proper restrictions are taken into account in the first summation; the fourth step follows from the fact that $1-\sum_{j=1}^{i}\pi(n(j),C) \leq 1-\sum_{j=1}^{i-1}\pi(n(j),C)$, while a strong inequality must be valid for at least one n(i); for the fifth step we used equation (13) and proposition 2 again; the last step should be obvious. This proves the second part of proposition 3. □

Appendix 2 Some further information about the data

As mentioned in the main text, we used the Dutch Housing Heeds Survey (WoningBehoefte Onderzoek, WBO) of 1981 (see DGVH [1982a]). A household was considered to be searching when it indicated to be desiring to move to another dwelling within a period of a year. We concentrated on the rented sector of the housing market and therefore restricted our attention to households currently occupying a rented dwelling who were searching for another rented dwelling.

The searching households usually indicated to what kind of dwelling they were planning to move : whether it should be a rented or an owner occupied dwelling, single family dwelling or apartment, the desired number of rooms and the rent one was willing to pay. It has been assumed that this information was based on up-to-date information about the situation on the housing market in the region concerned. It allowed us to make the classification of the stock of rented dwellings given in table A2.1, which was also used in Rouwendal [1989, chapter 10]. This is the classification that has been used in the text.

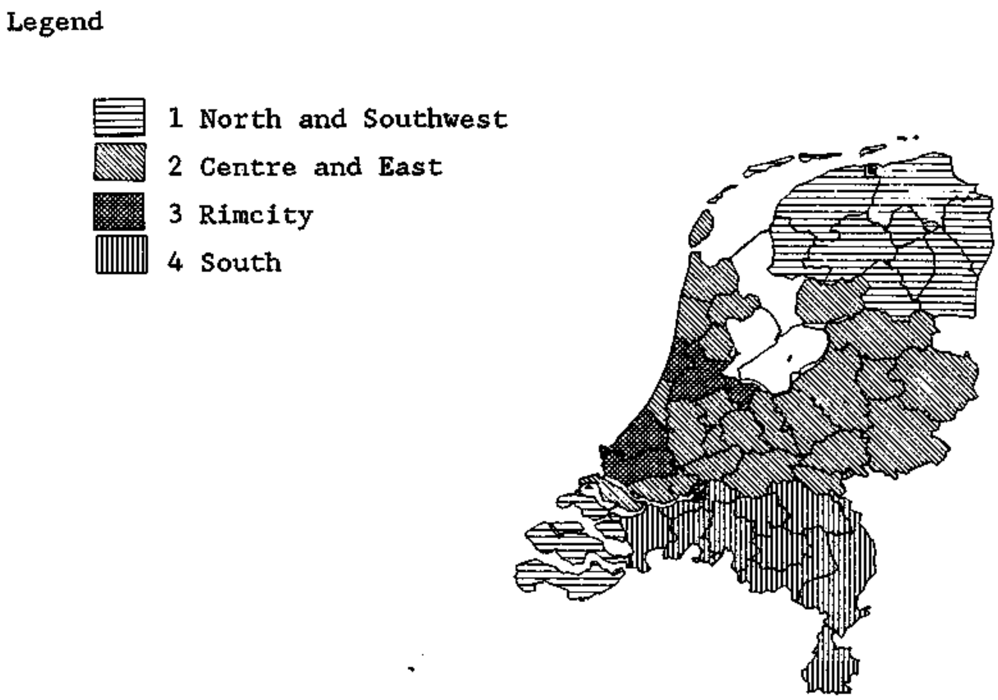
Table A2.1 Classification of rented dwellings

number	single fam./apartment	number of rooms	rent*
1	single family dwelling	1-3	< 250
2	,,	1-3	250-450
3	,,	1-3	> 450
4	,,	4	< 250
5	,,	4	250-450
6	,,	4	> 450
7	,,	≥ 5	< 450
8	,,	≥ 5	> 450
9	apartment	1,2	< 250
10	,,	1,2	> 250
11	,,	3	< 250
12	,,	3	250-450
13	,,	3	> 450
14	,,	≥ 4	< 250
15	,,	≥ 4	250-450
16	,,	≥ 4	> 450

* Dutch guilders per month

We have, furthermore, restricted ourselves to households who were willing to make a move within a short distance of their present location (i.e. we did not want to analyse migration). For this purpose we used the 'housing market areas', developed by the Dutch Ministry of Housing (see DGVH [1982b]). Since these area's were too small to use them as separate entities in our analysis, we aggregated them to the four regions used in the text. The aggregation that has been used was proposed by Scholten [1988, chapter 4], and is based on a clusteranalysis of mobility on the housing market. The basic area's and the four regions are pictured in figure A2.1.

The study of Scholten [1988] also contains valuable information about the stability of the patterns of mobility on the Dutch housing market in the period 1977-1981.



Figuur 1 Housing market areas and the four regions

Appendix 3 Estimation of the numbers of leavers

In order to estimate the numbers of households leaving the populations of searching households, we used the information about the realized duration of search contained in the Housing Needs Survey [WBO, 1981]. We used the simple assumption that the rates of inflow and outflow for the populations of searchers are constant over time. The constant rate of outflow will be denoted as q and is specific for the region and for the dwelling type which is the first best choice. The value of q can be estimated for each population of searching households in each of the four regions (see Rouwendal, 1989). The implied probability that a household looking for another dwelling will be searching for a period at least equal to t units of time equals $1\text{-}\exp(q.t)$. However, there appeared to be considerable differences between the rates of outflow of searchers with a short and a long realized duration of search. For this reason we estimated different models for those with a realized duration of search of more and less than a year. Although the differences in the estimates of q were in general large and significant, the predicted numbers of leavers remainde much the same. The correlation coefficients between both amounted to .99 for all four regions. In the text we used the figures of the model in which both groups of searchers were distinguished. For further details we refer to Rouwendal (1990).

1988-1	H. Visser	Austrian thinking on international economics
1988-2	A.H.Q.M. Merkies T. van der Meer	Theoretical foundations for the 3-C model
1988-3	H.J. Bierens J. Hartog	Nonlinear regression with discrete explanatory variables, with an application to the earnings function
1988-4	N.M. van Dijk	On Jackson's product form with 'jump-over' blocking
1988-5	N.M. van Dijk M. Rumsewicz	Networks of queues with service anticipating routing
1988-6	H. Linneman C.P. van Beers	Commodity Composition of Trade in Manufactures and South-South Trade Potential
1988-7	N.M. van Dijk	A LCFS finite buffer model with batch input and non-exponential sevices
1988-8	J.C.W. van Ommeren	Simple approximations for the batch-arrival $M^X/G/1$ queue
1988-9	H.C. Tijms	Algorithms and approximations for bath-arrival queues
1988-10	J.P. de Groot H. Clemens	Export Agriculture and Labour Market in Nicaragua
1988-11	H. Verbruggen J. Wuijts	Patterns of South-South trade in manufactures
1988-12	H.C. Tijms J.C.W. van Ommeren	Asymptotic analysis for buffer behaviour in communication systems
1988-13	N.M. van Dijk E. Smeitink	A non-exponential queueing system with batch servicing
1988-14	J. Rouwendal	Existence and uniqueness of stochastic price equilibria in heterogeneous markets
1988-15	H. Verbruggen	GSTP, the structure of protection and South-South trade in manufactures
1988-16	Mevr. H. Weijland Mevr. R. Herweijer J. de Groot	Female participation in agriculture in the Dominican Republic
1988-17	N.M. van Dijk	Product Forms for Random Access Schemes
1988-18	A.H.Q.M. Merkies I.J. Steyn	Adaptive Forecasting with Hyperfilters
1988-19	J. Rouwendal	Specification and Estimation of a Logit Model for Housing Choice in the Netherlands
1988-20	J.C.W. van Ommeren R.D. Nobel	An elementary proof of a basic result for the GI/G/1 queue
1988-21	H. Kool	A Note on Consistent Estimation of Heterostedastic and Autocorrelated Covariance Matrices
1988-22	C.P.J. Burger	Risk Aversion and the Family Farm
1988-23	N. van Dijk I.F. Akyildiz	Networks with mixed processor sharing parallel queues and common pools
1988-24	D.J.F. Kamann P. Nijkamp	Technogenesis: Incubation and Diffusion
1988-25	P. Nijkamp L. van Wissen A. Rima	A Household Life Cycle Model For the Housing Market
1988-26	P. Nijkamp M. Sonis	Qualitative Impact Analysis For Dynamic Spatial Systems
1988-27	R. Janssen P. Nijkamp	Interactive Multicriteria Decision Support For Environmental Management
1988-28	J. Rouwendal	Stochastic Market Equilibria With Rationing and Limited Price Flexibility
1988-29	P. Nijkamp A. Reggiani	Theory of Chaos in a Space-Time Perspective
1988-30	P. Nijkamp J. Poot J. Rouwendal	R & D Policy in Space and Time
1988-31	P. Nijkamp F. Soeteman	Dynamics in Land Use Patterns Socio-Economic and Environmental Aspects of the Second Agricultural Land Use Revolution
1988-32	J. Rouwendal P. Nijkamp	Endogenous Production of R & D and Stable Economic Development
1988-33	J.A. Hartog E. Hinloopen P. Nijkamp	Multicriteria Methoden: Een gevoeligheidsanalyse aan de hand van de vestigingsplaatsproblematiek van kerncentrales
1988-34	R. van der Mark P. Nijkamp	The Development Potential of High Tech Forms in Backward Areas - A Case study for the Northern Part of The Netherlands
1988-35	E.R.K. Spoor J.W.B. Vermeulen	Principes en gebruik van Envisage
1988-36	C. Gorter P. Nijkamp P. Rietveld	The Duration of Unemployment: Stocks and Flows on Regional Labour Markets in the Netherlands
1988-37	M. Hofkes	Parametrization of simplicial algorithms with an application to an empirical general equilibrium model
1988-38	J. van Daal A.H.Q.M. Merkies	A Note on the Quadratic Expenditure Model